

⑨ 日本国特許庁 (JP)

⑩ 特許出願公開

⑪ 公開特許公報 (A) 平4-144485

⑫ Int. Cl. 5

H 04 N 1/413  
G 06 F 15/66  
H 04 N 1/41  
1/415

識別記号

3 3 0  
3 3 0  
3 3 0

庁内整理番号

D  
J  
B  
C  
B

⑬ 公開 平成4年(1992)5月18日

審査請求 未請求 請求項の数 4 (全10頁)

⑭ 発明の名称 画像処理装置

⑮ 特願 平2-266675

⑯ 出願 平2(1990)10月5日

⑰ 発明者 三宅 英太 神奈川県海老名市本郷2274番地 富士ゼロックス株式会社  
海老名事業所内

⑰ 発明者 南 敏 東京都東大和市清水3丁目86番地13

⑰ 発明者 中村 納 東京都葛飾区堀切3-42-6

⑯ 出願人 富士ゼロックス株式会社 東京都港区赤坂3丁目3番5号

⑰ 代理人 弁理士 岩上 昇一 外2名

明細書

入力画像を一定間隔で代表値をとるサンプリング手段と、

サンプリングの結果から入力画像を曲面近似する近似手段と、

入力画像と前記近似手段の近似結果とを比較して近似誤差を計算する計算手段と、

この計算手段の計算結果が所定の閾値より小さい場合はそのまま符号化し、大きい場合は付加情報と付加して符号化する符号化手段とを有することを特徴とする画像処理装置。

(1) 請求項(2)に記載された画像処理装置により符号化された結果を復号する装置であって、符号化されたデータと付加情報と曲面近似情報とに分類する手段と、前記付加情報がある場合はこの付加情報を用いて復元された画像を補正する手段と

を有することを特徴とする画像処理装置。

(2) 多値画像を符号化する画像処理装置であって、

1. 発明の名称

画像処理装置

2. 特許請求の範囲

(1) 多値画像を符号化する画像処理装置であって、

入力画像を文字・図形等からなる文字領域と写真等の中間領域との二つに分類する分類手段と、この分類手段の分類の結果、文字領域に分類された領域に対し入力画像を  $N_1 \times N_2$  ( $N_1, N_2$  は正整数) 画素のブロックに分割するブロック化手段と、

このブロック化手段によるブロック化された各ブロックを符号化する第1の符号化手段と、

前記分類手段の分類の結果、中間領域に分類された領域に対し符号化する第2の符号化手段とを有することを特徴とする画像処理装置。

(2) 多値画像を符号化する画像処理装置であって、

PTO 2000-3063

S.T.I.C. Translations Branch

きくすることを特徴とする請求項(2)記載の画像処理装置。

### 3. 発明の詳細な説明

#### 【産業上の利用分野】

本発明は、文字領域および中間調領域を含む多値画像を符号化し、あるいはそのように符号化された符号を復号する処理を行う画像処理装置に関するものである。

#### 【従来の技術】

画像は直感的に全体を把握することのできる情報であり、画像を用いたコミュニケーションは情報伝達機能の極めて優れた手段として、広く一般に普及している。しかしながら、画像情報は一般に情報量が極めて多いため、これを原画像情報の形のままでは通信網を通して遠隔地へ伝送したり各種電子媒体に蓄積するとコストおよび処理時間が大きな制約となってしまう。このために画像情報を効率的に圧縮し、伝送または蓄積するための高能率符号化方式が必要になってくる。

高能率符号化方式としては、予測符号化方式、

-3-

を  $N_1 \times N_2$  ( $N_1, N_2$  は正整数) 画素のブロックに分割するブロック化手段(第1図12)と、このブロック化手段によるブロック化された各ブロックを符号化する第1の符号化手段(第1図13、第3図)と、分類手段の分類の結果、中間調領域に分類された領域に対し符号化する第2の符号化手段(第1図14、第5図)とを有する。

また、本発明の多値画像を符号化する画像処理装置は、中間調領域を符号化するために、入力画像を一定間隔で代表値とするサンプリング手段(第5図51)と、サンプリングの結果から入力画像を曲面近似する近似手段(第5図53)と、入力画像と前記近似手段の近似結果とを比較して近似誤差を計算する計算手段(第5図54)と、この計算手段の計算結果が所定の閾値より小さい場合はそのまま符号化し、大きい場合は付加情報を付加して符号化する符号化手段(第5図55～58)とを有する。

また、本発明の画像処理装置により符号化された結果を復号する装置は、符号化されたデータを

ブロック符号化方式(BTC:Block Truncation Coding)、直行変換符号化方式、ベクトル量子化方式、エントロピー符号化方式などが知られている。

#### 【発明が解決しようとする課題】

このように様々な高能率符号化方式が提案されているが、対象とする文書画像の性質によりそれぞれ一長一短があり、文字、線画、中間調画像などが混在する文書の画像を符号化するには十分ではなかった。

本発明は、画像情報を高い画品質を保持し、しかも高度に圧縮することができ、処理時間も短い符号化／復号化を行うことのできる画像処理装置を提供することを目的とするものである。

#### 【課題を解決するための手段】

本発明は、多値画像を符号化する画像処理装置であって、入力画像を文字・図形等からなる文字領域と写真等の中間調領域との二つに分類する分類手段(第1図11)と、この分類手段の分類の結果、文字領域に分類された領域に対し入力画像

-4-

附加情報と曲面近似情報とに分類する手段(第1図111)と、前記附加情報がある場合はこの附加情報を用いて復元された画像を補正する手段(第11図114～117)とを有する。

また、本発明の一態様によれば、近似手段(第5図53)は曲面の大きさを画像の濃度変化に応じて変化させ、変化が激しい場合は曲面を小さくし、変化が少ない單調な場合は曲面を大きくする構成とができる。

#### 【作用】

入力画像を領域分類手段により文字領域と中間調領域に分離し、文字領域はブロック化手段によりブロックに分割する。第1の符号化手段は文字領域の統計的性質に適した符号化方式により各ブロックの符号化を行う。また、第2の符号化手段は中間調領域の統計的性質に適した符号化方式により符号化を行う。本発明はこのように画像を領域に分け、それぞれの領域に適した符号化方式を用いて符号化するので、高い画品質で高度の圧縮を行うことができ、処理時間も短縮することができる。

きる。

また、本発明の符号化においては、サンプリング手段によりサンプリングし、サンプル点の画素濃度を符号化するので、符号量を高度に圧縮することができる。さらに近似手段が曲面近似によりサンプル点から補間した画像と入力画像との差（近似誤差）が既定の閾値より大きいときは、符号化手段は近似誤差を付加情報として符号化し、前記サンプル点の画素濃度の符号化情報に付加する。従って、復号に際して付加情報により誤差を補正することができるので、高品質で符号化復号化を行うことができる。

また、本発明はサンプル点を画像の性質に応じて濃度変化の激しい領域には細かいサンプリング、単純な領域には粗いサンプリングを行うことにより、曲面近似のパッチサイズを可変にすれば、効率的な符号化が可能となる。

#### 【実施例】

第2図は本発明をデジタル複写機に適用した画像処理装置の概略の構成を示すもので、イメージ

スキナ等の画像入力部21、画像の変換や組合せ等の他の画像の処理を行う画像処理部22と、プリンタ等の画像出力部23と、画像を圧縮して効率化符号化を行い、あるいはその逆の処理による復号化を行う符号化／復合化部(CODEC)24と、符号化された画像を記憶する記憶部25とからなっている。

第1図は本発明の符号化を行う実施例の構成を示すもので、入力画像を文字・图形等からなる文字領域と写真等の中間調領域との二つに分類する領域分類部11と、この領域分類部11の分類の結果、文字領域に分類された領域に対し入力画像を $N_1 \times N_2$ ( $N_1, N_2$ は正整数)画素のブロックに分割するブロック化部12と、このブロック化部12によりブロック化された各ブロックを符号化する文字領域符号化部13と、領域分類部11の分類の結果、中間調領域に分類された領域に対し符号化する中間調領域符号化部14とを備えている。領域分類部11とブロック化部12は第2図の画像処理部22の機能に属し、文字領域符号

-7-

化部13と中間調領域符号化部14は第2図の符号化／復号化部24の機能に属している。

画像入力部21によって入力された文書画像は領域分類部11により、文章、グラフ、線図等からなる文字領域と、網点やディザを用いて淡談を表す図および写真等を含む中間調領域に分離する。その分離は画像の性質を利用して自動的に行っても、あるいはオペレーターが指示するようにしてもよい。

ブロック化部301は、文字領域を $N_1 \times N_2$ 画素のブロックに分割する。本実施例では $4 \times 4$ 画素ブロックとした。

各ブロックは文字領域符号化部13により高能率符号化され、記憶部25により記憶される。

ここで、文字領域符号化部13について詳細に説明する。

第3図は第1図における文字領域符号化部13を詳細に示す図である。

文字領域符号化部13は、ブロック化部301(第2図12)により切り出された文字領域のブ

-8-

ロックを記憶するブロックメモリ302と、ブロック内の平均濃度値を算出する平均濃度値算出部303と、ブロック内の標準偏差を算出する標準偏差算出部304と、算出した平均濃度値および標準偏差値を基に、画像の濃度変化を調べ、画像の性質を判定するラベル判定部305と、ラベル判定部305の判定の結果、そのブロックが濃度変化の非常に緩やかな部分(ラベル①〔背景部〕)に属するとき、そのブロックの符号化に用いる1階調ブロック近似符号化部(1階調BTC)306と、そのブロックが濃度変化の比較的緩やかな(ラベル②〔卓背景部〕)に属するとき、そのブロックの符号化に用いる2階調ブロック近似符号化部(2階調BTC)307と、そのブロックが濃度変化の激しい部分(ラベル③〔文字部〕)に属するとき、線形量子化する線形量子化部308と、1階調BTC306、2階調BTC307または線形量子化部308からの階調情報を圧縮符号化するハフマン符号化部309と、1階調BTC306、2階調BTC307および線形量子化

部 308 からの情報をまとめて圧縮符号化する M R 符号化部 310 と、符号を組み立てる符号組立部 311 からなっている。

次に、文字領域符号化部の動作について説明する。

ブロックメモリ 302 に記憶された文字領域のブロックは、平均濃度値算出部 303 で平均濃度値  $A_{ve}$  算出し、標準偏差算出部 304 で標準偏差  $S_d$  を算出する。ラベル判定部 305 は平均濃度値  $A_{ve}$ 、標準偏差  $S_d$ 、および適宜に定めた閾値  $T_{h1}, T_{h2}$  を基に、次のように、ラベルの判定を行う。

ラベル① (背景部) :  $A_{ve} > T_{h1}$

ラベル② (準背景部) :  $A_{ve} \geq T_{h1}$  かつ  
 $S_d < T_{h2}$

ラベル③ (文字部) :  $A_{ve} \leq T_{h1}$  かつ  
 $S_d \geq T_{h2}$

ラベル判定部 305 でラベル①と判定されたときは、そのブロックは濃度変化が緩やかで平坦な部分である。この統計的性質からブロック内を 1 階調で近似する。そのため 1 階調ブロック近似符号化部 306 を用い、ブロック内の平均濃度値をブロックの階調とする。

ラベル判定部 305 でラベル②と判定されたときは、そのブロックは濃度変化が平坦ではないが緩やかな部分である。この統計的性質からブロック内を 2 階調で近似する。そのため 2 階調ブロック近似符号化部 307 を用い、2つの濃度代表値からなる階調情報  $X_1, X_2$  と、各画素がどちらの代表値で近似されるのかを示す分解能情報  $S_{1,2}$  を求める。

ラベル判定部 305 でラベル③と判定されたときは、そのブロックは濃度変化が激しい部分である。この統計的性質に適した処理として線形量子化を用いる。すなわち、線形量子化部 308 により、背景部分について 1 レベルを設定し、それ以外の部分を 15 等分してレベルを設定した 16 階調の量子化レベルにより、線形量子化する。

ハフマン符号化部 309 では、1 階調ブロック近似符号化部 306、2 階調ブロック近似符号化

部 307 および線形量子化部 308 の出力する階調情報を可変長符号割り当てによる圧縮をするハフマン符号化を行う。

また、MMR 符号化部 310 は、ラベル①の 1 階調ブロック近似符号化部 306 からの出力に対しては各画素のすべてに 0 を割当て、ラベル②の 2 階調ブロック近似符号化部 307 により得られた分解能情報を基にし、ラベル③の線形量子化部 308 からの出力に対しては各画素のすべてに 1 を割当てまとめて MMR 符号化する。

符号組立部 311 はハフマン符号化部 309 および MMR 符号化部 310 の出力を基に符号化出力を生成する。第 4 図は一つの文字領域を記憶部 25 に蓄積するために組み立てる符号構成を示すもので、以下の情報を符号化する。

ヘッダ情報 H : 文字領域の画像サイズや領域の位置

ラベル① : 階調情報 G

ラベル② : 階調情報 G および分解能情報 P

ラベル③ : 階調情報 G

以上のようにして符号化された文字画像は記憶部 25 に記憶される。

次に、文字領域について符号化され記憶された情報の復元は、階調情報、分解能情報から各ブロックの濃度分布を再生することにより行う。ラベル①とラベル②のブロックに関しては各ブロックを 1 階調、2 階調（代表値）で再生する。ラベル③に関しては、16 個の代表値によってそのブロックの各画素の濃度値を再現する。

次に、中間調領域の処理について説明する。

第 5 図は、第 1 図の中間調領域符号化部 14 の構成の一例を示すもので、サンプリング部 51、予測符号化部 52、曲面近似部 53、近似誤差計算部 54、非線形量子化部 55、ハフマン符号化部 56、モード判定部 57 および符号組立部 58 からなっている。

この例における符号化のアルゴリズムは、まず、縦、横、斜め方向の濃度値の相図を利用するため、サンプリング部 51 により五の目状にサンプリングを行う。このサンプリング位置は固定している

ので、サンプル点の位置を表す情報は不要である。

次に、サンプル点の濃度値を、隣接するサンプル点における濃度値の分散が小さいことから、符号量の削減のために予測符号化部 52 により予測符号化を行う。

そして、隣接するサンプル点 4 点で囲まれた近似対象領域を双三次パッチを用いた曲面近似部 53 により近似し、領域内の原画像と近似画像との平均 2 乗誤差を近似誤差計算部 54 により各領域ごとに求める。

モード判定部 57 は平均 2 乗誤差が閾値以上なら領域内の各画素の近似誤差（付加情報）を非線型量子化部 55 およびハフマン符号化部 56 により圧縮符号化する。

以下、各処理についてさらに詳細に説明する。

サンプリング部 51 は縦、横、斜め方向の濃度の相間を利用するため、第 6 図のように五の目状にサンプリングを行う。サンプル点の間隔は近似する際の濃度変化が激しい部分でのボケを最小

限に軽減するために 4 画素とする。また、サンプル点の濃度値は第 6 図の斜線で示す 4 画素の平均濃度値とする。サンプル点は画素ではなく、画素と画素の境界に位置する。

予測符号化部 52 は、第 7 図に示す柱目サンプル点の周辺に隣接する参照サンプル点 4 点の平均値を予測値とし、柱目サンプル点の濃度値からこの予測値を引いて予測誤差を求める。この予測誤差に対して、非線型量子化を施し、18 個の代表値で近似する。次に、各代表値にはハフマン符号を割り当てて符号化する。

曲面近似部 53 は、サンプル点 4 点でか込まれた領域（4 頂点が頂点となる複数の領域、以下、近似対象領域と呼ぶ）を、双三次パッチを適用し近似する処理を行う。双三次パッチでは第 8 図に示す 16 個のサンプル点を参照して近似対象領域を内挿する。

近似誤差計算部 54 は、近似対象領域内において、どの程度誤差があるかを調べるために、第 9 図に示すように領域内の原画像と近似画像の平均

2 乗誤差 MSE を各領域ごとに求める。

モード判定部 57 は、平均 2 乗誤差 MSE が閾値 TH 以上ならモード 1（付加情報が必要である領域）、閾値 TH 未満ならモード 2（付加情報が必要でない領域）と判定する。ここで近似対象領域がどちらのモードに属するかを示すモード情報 MD を作成するモード情報は 0, 1 の 2 値画像であるから、符号化する際は MMR 符号化を採用する。

モード判定部 57 の判定がモード 1 であるとき、近似誤差を非線型量子化部 55 およびハフマン符号化部 56 により符号化する。この実施例では非線形量子化部 55 は 11 個の代表値で近似する。そしてハフマン符号化部 56 では各代表値にハフマン符号を割り当てて符号化する。

符号組立部 58 は、サンプル点の予測誤差 S と、モード情報 M と、近似対象領域内の各画素の近似誤差、つまり付加情報 F を第 10 図のような符号構成で組み立てて符号化出力とする。符号化出力の最初にはヘッダ情報 H を挿入する。ヘッダ情報

H は中間調領域の画像を符号化する際の領域の位置、大きさ等の情報である。なお、付加情報はモード 1 の場合のみ必要である。

以上のように、符号化した中間調領域の画像の復号は、例えば第 11 図に示すような構成とする復号化部により実行することができる。第 10 図のような符号構成の符号情報から、情報分離部 111 で予測誤差情報 S、付加情報 F、モード情報 M 等を分離し、それぞれの復号部 112、114、117 で復号する。予測符号化符号の復号部 112 から出力されるサンプル点の濃度情報を基づき曲面近似部 113 は双三次パッチにより内挿し、各画素の濃度値を求める。モード情報がモード 2 のときは選択部 116 により曲面近似部 113 の出力をそのまま復号出力として選択出力する。モード情報がモード 1 のときは補正部 115 により曲面近似部 113 の出力を付加情報復号部 114 で復号した近似誤差値で補正し、選択部 116 では補正部 116 の出力を復号出力として選択出力する。

域には細かいサンプリング、單調な領域には粗いサンプリングを行えば効率的な符号化が可能である。

第2の実施例が第5図に示す第1の実施例と構成（機能）上異なるところは、主としてサンプリング部51におけるサンプル点の選び方がパッチサイズに基づくことであり、その他の構成はほぼ同じである。

パッチは第12図に示すように4つのサンプル点を頂点とする菱形の領域で、画素の並びに対し $45^{\circ}$ 傾いた正方形である。また、サンプル点は第1の実施例と同様に4画素の中心に仮想的にとることとし、この4画素の平均値をサンプル点の濃度値とする。

パッチサイズは、最小のときでmが2であり、濃度変化に応じて、第12図に示すようにパッチの最上部に位置する点を固定した状態でパッチを約縦状に拡大し、nを0から7の範囲で決定する。

その決定の手順は次のとおりである。

I) 第13図に示すように、最小パッチサイ

-20-

存在するとは限らない、②1つも点が存在しない場合がある。

このことを考慮して各々の問題点に対して次のように処理する。

①の問題点は、存在している点だけを使い平均値を求める。

②の問題点に関しては、パッチの最上部に位置する頂点の値を予測値として使用する。画面最上段の場合最上部に位置する頂点も値が存在しないので、使用可能な濃度範囲の中央値を予測値とする。

符号構成は、第10図の構成にさらにパッチの最上部の頂点の座標およびパッチサイズの情報を加わる。

#### 【発明の効果】

本発明によれば、対象とする文書画像を中間領域と文字領域との2領域に分離し、各領域の統計的な性質に適した符号化方式を用いるようにしたので、高い画品質を保ちながら、高度の圧縮を行うことができる。

以上の実施例（第1の実施例）において、中間領域のパッチサイズを一定とした場合すなわちサンプル点を固定的な一定間隔のものとした場合を説明したが、サンプル点を画像の性質に応じて可変とし、パッチサイズを変えるように構成することもできる。

以下、第2の実施例としてそのパッチサイズ（従ってサンプル点）の可変の方式について説明する。

この方式では、符号化する画像を多数の菱形（正方形を傾斜させた形）、すなわちパッチに分け、各々のパッチを構成するのに必要な点をサンプル点として抽出し、このサンプル点の濃度情報を蓄積する。復号化においては、このサンプル点の濃度情報に基づいて、パッチ内の濃度値を保管し、すべてのパッチについてこれを行うことで画像を復元する。パッチサイズの決め方は、中間領域には、物体表面などの濃度変化の激しい部分と背景のように濃度分布の單調な領域が混在していることを基礎としている。濃度変化の激しい領

-19-

域（n=0）状態の濃度勾配情報を基準情報として求める。

II) n=1のパッチを最小パッチサイズに分割し、各々の最小パッチについて基準情報と同様に濃度勾配を求める。

III) 濃度勾配の絶対値と基準情報の絶対値の差の絶対値が閾値（TH1）以上ならn=0をパッチサイズとする。閾値以下ならn=2として上と同様な処理を行う。ただし、既にパッチが当っている最小パッチが含まれている場合は、濃度勾配に関係なくn=0とする。

IV) nを0から7まで変化させてパッチサイズの大きさを決定する。

パッチサイズが決るとサンプル点を抽出することができる。各サンプル点に対しては予測符号化を行う。予測符号化は第1の実施例において述べたのと同様にして行うことができる。ただ、第1の実施例とは異なり、パッチサイズを変更しながらサブサンプリングを行うため、二つの問題が生ずる。すなわち、①予測に用いる4点が必ずしも

第 5 図は第 1 図における中間調領域符号化部の構成例を示す図である。

第 6 図は固定密度サンプリングのサンプル点を示す図である。

第 7 図は予測符号化の参照サンプル点を示す図である。

第 8 図は双三次パッチの座標形を示す図である。

第 9 図は領域内の原画像と近似画像の平均 2 乗誤差を説明するための図である。

第 10 図は中間調領域の符号化の符号構成を示す図である。

第 11 図は中間調領域の復号化を行うための構成を示す図である。

第 12 図はパッチサイズの決定の仕方を説明するための図である。

第 13 図は傾度勾配を説明するための図である。

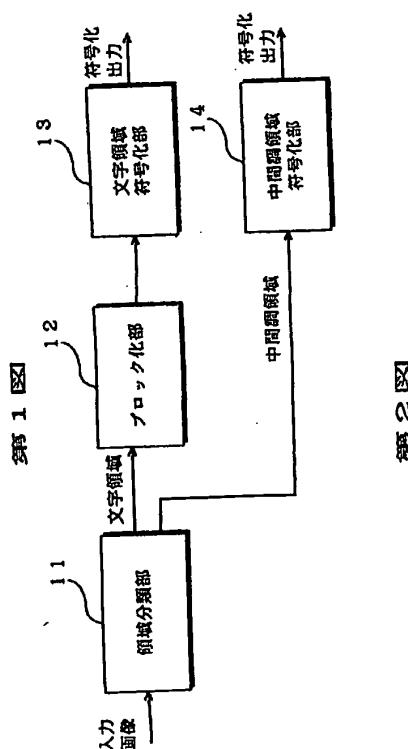
特許出願人 富士ゼロックス株式会社

代理人弁理士 岩上昇一

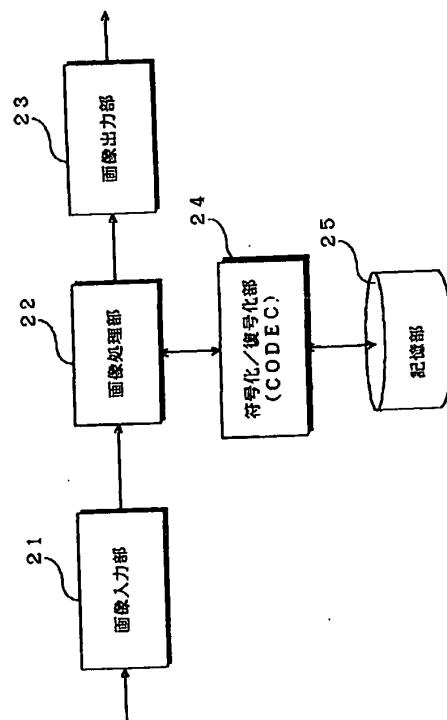
(外 2 名)

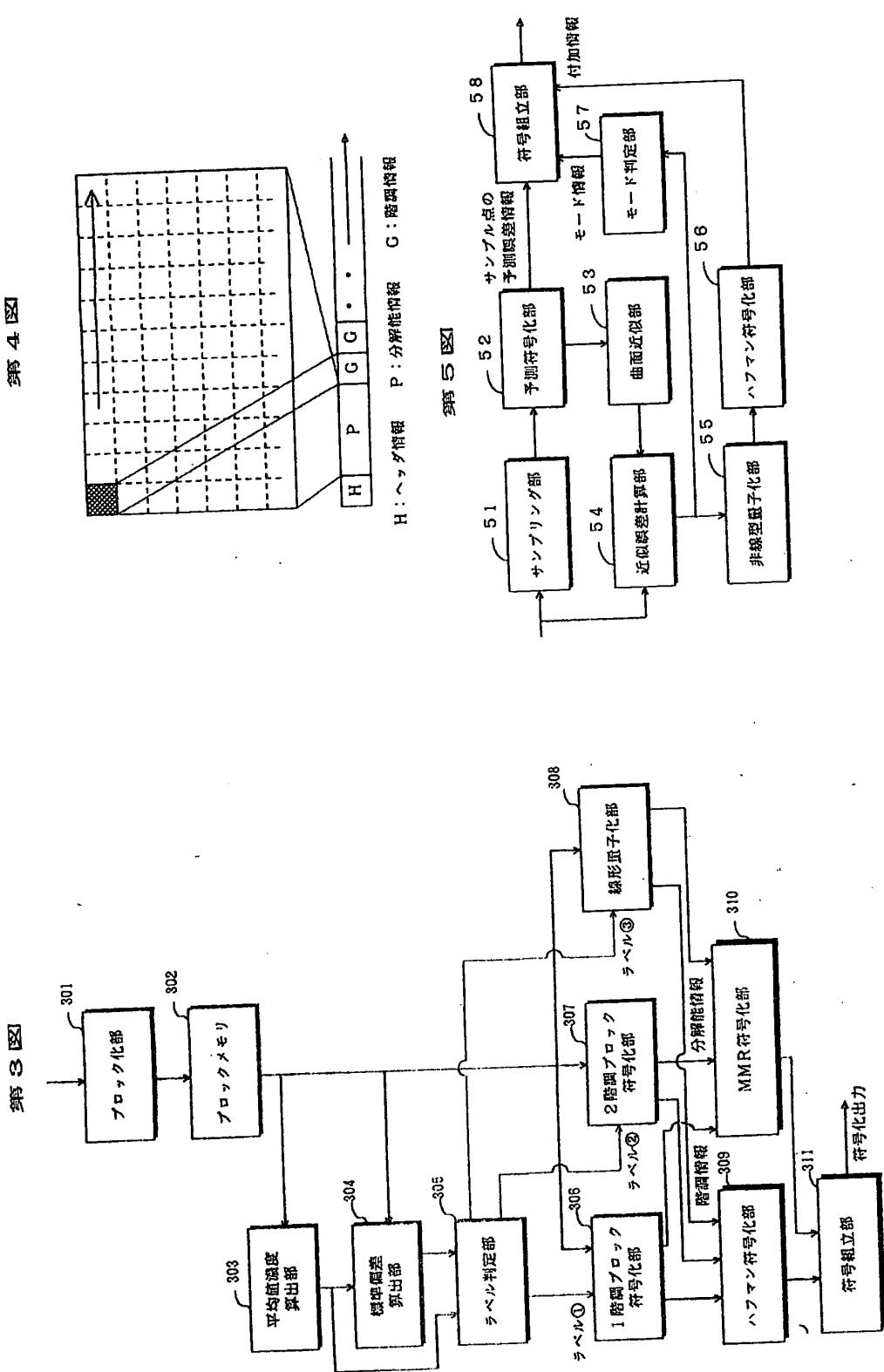
-23-

-24-

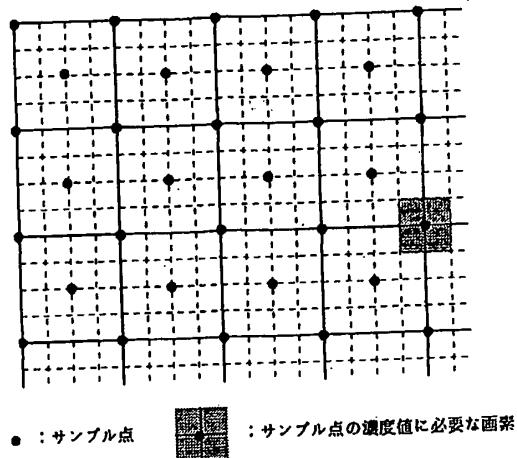


第 2 図

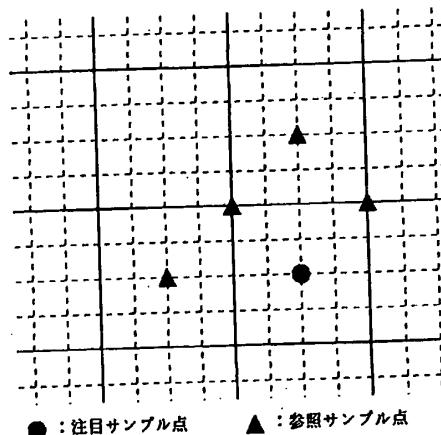
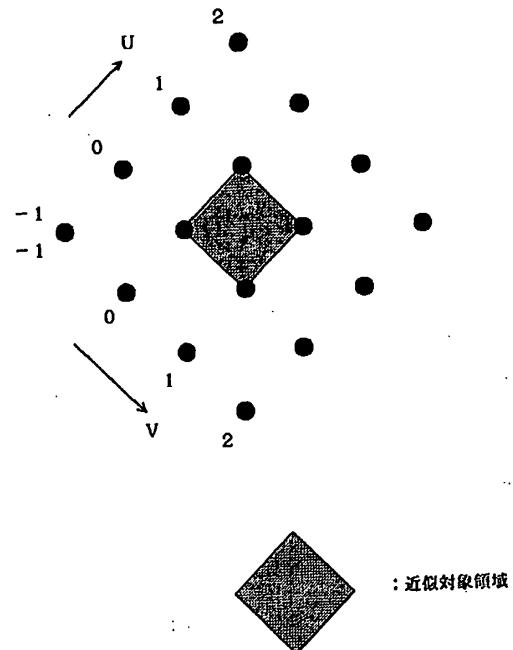




第6図 固定密度サンプル

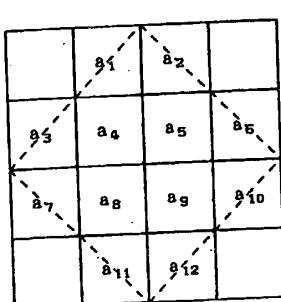


第7図 予測誤差の算出

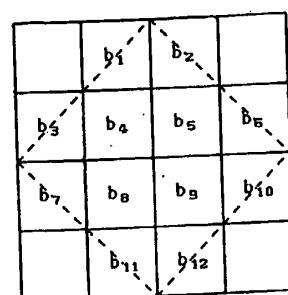
第8図  
双3次バッチの座標系

第9図

(a) 原画像

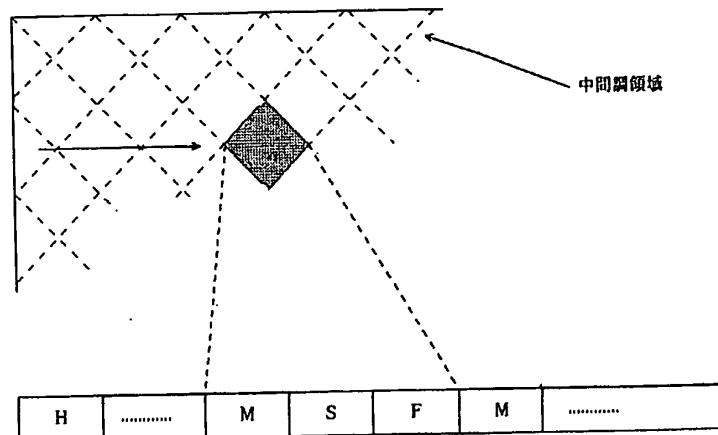


(b) 近似画像



$$\text{平均2乗誤差MSE} = \frac{1}{12} \sum_{i=1}^{12} (b_i - a_i)^2$$

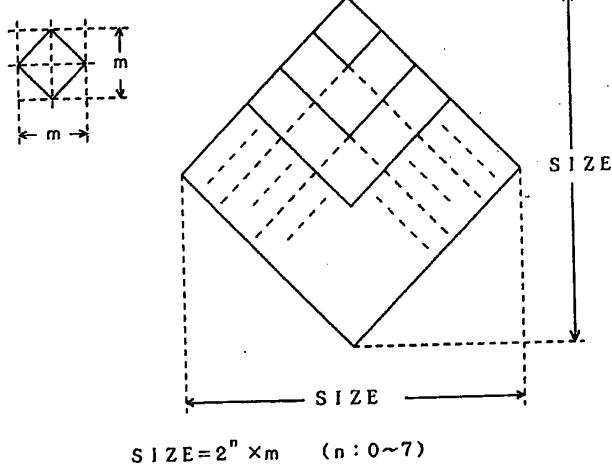
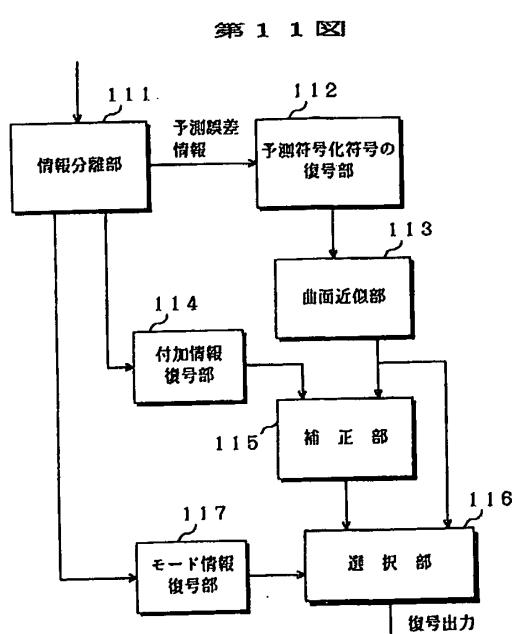
第10図 符号樹成



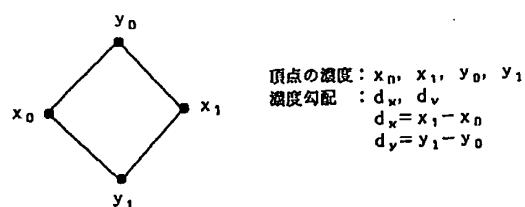
H : ヘッジ情報      S : サンプル点の予測誤差  
M : モード情報      F : 付加情報 (モード1のみ)

第12図

パッチの大きさ



第13図



PTO 00-3063

Japanese Kokai Patent Application  
No. Hei 4 [1992]-144485

IMAGE PROCESSING DEVICE

Eita Miyake, et al.

UNITED STATES PATENT AND TRADEMARK OFFICE  
WASHINGTON, D.C. JUNE, 2000  
TRANSLATED BY THE RALPH MCELROY TRANSLATION COMPANY

Code: PTO 00-3063

JAPANESE PATENT OFFICE  
PATENT JOURNAL (A)  
KOKAI PATENT APPLICATION NO. HEI 4 [1992]-144485

Int. Cl.<sup>5</sup>: H 04 N 1/413  
G 06 F 15/66

H 04 N 1/41  
1/415

Sequence Nos. for Office Use: 8839-5C  
8420-5L

Filing No.: Hei 2[1990]-266675

Filing Date: October 5, 1990

Publication Date: May 18, 1992

No. of Claims: 4 (Total of 10 pages)

Examination Request: Not filed

IMAGE PROCESSING DEVICE  
[Gozo shori sochi]

Inventors: Eita Miyake, et al.

Applicant: Fuji Xerox, KK

[There are no amendments to this patent.]

/1\*

Claims

1. Image processing device that codes multiple value images, characterized by the fact that it comprises a classification means that classifies the input image into two regions; i.e., the character region consisting of letters, figures, etc., and the halftone region of photographs, etc.,

---

\* [Editor's note: Pagination in the right margin represents pagination in the original foreign text.]

a block-making means that divides the input image classified into the character region as a result of the classification into blocks consisting of  $N_1 \times N_2$  ( $N_1$  and  $N_2$  are positive integers) elements,

a first coding means that codes each of the blocks that are made into blocks by this block-making means,

and a second coding means that codes the region that has been classified into the halftone region as a result of the classification by the aforementioned classification means.

2. Image processing device that codes multiple value images, characterized by the fact that it comprises a sampling means that draws representative values at constant intervals of the input image,

an approximation means that approximates in curved surfaces the input image from the result of sampling,

a computation means that computes the approximation errors by comparing the input image and the result of approximation by the aforementioned approximation means,

and a coding means which codes directly when the result of computation by this computation means is smaller than a prescribed threshold value, but codes by adding additional information in case it is larger.

3. Device that decodes the result coded by the image processing device in accordance with Claim 2, characterized by the fact that it comprises a means to classify the coded data into additional information and curved-surface approximation information, and a means to correct the decoded image using this additional information in case the aforementioned additional information is present.

4. Image processing device in accordance with Claim 2 characterized by the fact that the approximation means changes the dimensions of the curved surface corresponding to the change in the concentration of the image, and that it reduces the curved surface when there are many changes and increases the curved surface in a monotone case where there are few changes.

/2

#### Detailed explanation of the invention

##### Industrial application field

The present invention pertains to an image processing device that codes multiple value images including character region and halftone region, or decodes images that are coded in that manner.

##### Prior art

Images are information that can assist in intuitively grasping the whole. Communication using images is widely popularized as an excellent means to transmit information. However,

because images generally contain an extremely large quantity of information, if they are directly transmitted to a remote location through communication network in the form of original image information or stored in various electronic media, the cost and the processing time become a significant restriction. Therefore, a high efficiency coding system to efficiently compress, transmit or store image information is needed.

Known high efficiency coding systems include the predictive coding system, the block coding system (BTC: Block Truncation Coding), the direct conversion coding system, the vector quantization system, entropy coding system, etc.

#### Problems to be solved by the invention

Although various high efficiency coding systems have been proposed as mentioned earlier, each has its advantages and disadvantages depending on the characteristics of the document image. None is good enough to code images of documents wherein letters, line drawings, halftone images, etc., are mixed.

The purpose of the present invention is to offer an image processing device that maintains the high picture quality of the image information, can perform advanced compression, and can code/decode in a short processing time.

#### Means to solve the problems

The present invention is an image processing device that codes multiple value images, and comprises a classification means (11 in Figure 1) that classifies the input image into two regions; i.e., the character region consisting of characters, figures, etc., and the halftone region of photographs, etc., a block-making means (12 in Figure 1) that divides the input image classified into the character region as a result of the classification by this classification means into blocks consisting of  $N_1 \times N_2$  ( $N_1$  and  $N_2$  are positive integers) elements, a first coding means (13 in Figure 1 and Figure 3) that codes each of the blocks that are made into blocks by this block-making means, and a second coding means (14 in Figure 1 and Figure 4) that codes the region that has been classified into the halftone region as a result of the classification by the aforementioned classification means.

Also, it is an image processing device characterized by the fact that it comprises a sampling means (61 in Figure 5) that draws representative values at constant intervals of the input image, an approximation means (53 in Figure 5) that approximates in curved surfaces the input image from the result of sampling, a computation means (54 in Figure 5) that computes the approximation errors by comparing the input image and the result of approximation by the aforementioned approximation means, and a coding means (55-58 in Figure 5), which codes

directly when the result of computation by this computation means is smaller than a prescribed threshold value, but codes by adding additional information in case it is larger.

Also, the device that decodes the results coded by the image processing device in accordance with the present invention comprises a means to classify (111 in Figure 11) the coded data into additional information and curved-surface approximation information, and a means to correct (114-117 in Figure 11) the decoded image using this additional information in case the aforementioned additional information is present.

Also, according to one mode of the present invention, the approximation means (53 in Figure 5) can have a constitution wherein it changes the dimensions of the curved surface corresponding to the change in the concentration of the image, so that it reduces the curved surface when there are a lot of changes and increases the curved surface in a monotone case where there are few changes.

#### Function

The input image is classified into the character region and the halftone region by the region classification means, and the character region is divided into blocks by the block-making means. The first coding means codes each block according to a coding system that is appropriate for the statistical characteristics of the character region. Also, the second coding means codes the halftone region according to a coding system that is appropriate for its statistical characteristics. In this manner, in the present invention the image is divided into regions and coded using the coding system that is appropriate to each of the regions. Therefore, advanced compression can be made while maintaining a high picture quality, and processing time can also be shortened.

Also, in the coding of the present invention, the sampling means takes samples and codes the picture element concentration at the sample point. Thus, the amount of coding can be compressed at an advanced level. Furthermore, when the difference (approximation difference) between the image the approximation means interpolated from the sample point by curved-surface approximation and the input image is larger than the prescribed threshold value, the coding means codes the approximation error as additional information and adds it to the coding information of the picture element concentration for the aforementioned sample point. Therefore, errors can be corrected by means of additional information at the time of decoding. Thus, decoding can be performed at a high quality.

Also, the present invention can vary the sampling depending on the characteristics of the image; i.e., fine sampling for regions with severe concentration changes and rough sampling for monotone regions. In this manner, the batch size of curved-surface approximation can be made variable, and efficient coding can be realized.

### Application examples

Figure 2 illustrates a schematic constitution of an image processing device wherein the present invention is applied to a digital copier. It comprises an image input part (21) such as an image scanner, an image processing part (22) that performs processing of images such as image conversion or editing, an image output part (23) such as a printer, a coding/decoding means (CODEC) (24) that performs efficiency coding by compressing the image or decoding by the inverse processing, and a memory (25) that stores the coded images.

Figure 1 illustrates the constitution of an application example that performs the coding of the present invention. It comprises a region classification means (11) that classifies the input image into the character region consisting of characters, figures, etc., and the halftone region such as photographs, a block-making means (12) that divides the input image of the region classified into the character region as a result of the classification by this region classification means (11) into blocks of  $N_1 \times N_2$  ( $N_1$  and  $N_2$  are positive integers) picture elements, a character region coding means (13) that codes each of the blocks that are made into blocks by this block-making means (12), and a halftone region coding means (14) that codes the region that has been classified into the halftone region. Both the region classification means (11) and the block-making means (12) belong in the function of the image processing part (22) in Figure 2, and both the character region coding means (13) and the halftone region coding means (14) belong in the function of the coding/decoding means (24) in Figure 2.

The document image input by the image input part (21) is separated by the region classification means (11) into the character region consisting of sentences, graphs, line drawings, etc., and the halftone region including figures that express shades using halftone dots or dithers, photographs, etc. The separation may be automatically made using the characteristics of the image or may be indicated by the operator.

The block-making means (301) divides the character region into blocks of  $N_1 \times N_2$  picture elements. In the present application example, blocks of  $4 \times 4$  picture elements were adopted.

Each block is high efficiency coded by the character region coding means (13), and stored by the memory (25).

Here we will explain the character region coding means (13) in detail.

Figure 3 illustrates the character region coding means (13) in Figure 1 in detail.

The character region coding means (13) comprises a block memory (302) that stores the blocks of character regions created by the block-making means (301) (12 in Figure 2), a mean concentration value computing means (303) that computes the mean concentration value within the block, a standard deviation computing means (304) that computes the standard deviation within the block, a label determination means (305) that examines the concentration changes of

the image and determines the characteristics of the image based on the computed mean concentration value and the standard deviation value, a one gradation block approximation coding means (one gradation BTC) (306), which is used to code the block when the block belongs in a part (label ① [background part]) where the concentration change is very mild as a result of the determination by the label determination means (305), a two gradation block approximation coding means (2 gradation BTC) (307), which is used to code the block when the block belongs in (label ② [quasi background part]) where the concentration change is relatively mild, a linear quantization part (308), which performs linear quantization when the block belongs in a part (label ③ [character part]) where the concentration change is severe, a Huffman coding means (309) that compression codes the gradation information from the one gradation BTC (306), two gradation BTC (307) or linear quantization part (308), an MMR coding means (310) that compression codes all information from the one gradation BTC (306), two gradation BTC (307) and linear quantization part (308) together, and a code assembly part (311) that assembles the codes.

Next, we will explain the action of the character region coding means.

The character region block stored in the block memory (302) has the mean concentration value computing means (303) compute the mean concentration value (Ave), and has the standard deviation computing means (304) compute the standard deviation (Sd). The label determination means (305) will determine the label as follows based on the mean concentration value (Ave), standard deviation (Sd) and appropriately determined threshold values (Th1) and (Th2).

Label ① (Background part) : Ave > Th1

Label ② (Quasi background part) : Ave  $\geq$  Th1, and  
Sd < Th2

Label ③ (Character part) : Ave  $\geq$  Th1, and  
Sd  $\geq$  Th2

When a block is determined to be label ① by the label determination means (305), the block is a flat section where the concentration change is mild. Based on this statistical characteristic, the inside of the block is approximated in one gradation. Therefore, using the one gradation block approximation coding means (306), the mean concentration value within the block is assumed to be the gradation of the block.

When a block is determined to be label ② by the label determination means (305), the block is a section where the concentration change is not flat but mild. Based on this statistical characteristic, the inside of the block is approximated in two gradations. Therefore, using the two gradation block approximation coding means (307), the gradation information ( $X_0$ ) and ( $X_1$ ) consisting of two representative values of concentration and the resolving power information ( $S_{11}$ ), which indicates which representative value approximates each picture element, is obtained.

When a block is determined to be label ③ by the label determination means (305), the block is a part where the concentration change is severe. As a process appropriate for this statistical characteristic, linear quantization is adopted. That is, the linear quantization part (308) performs linear quantization using 16 gradations of quantization levels obtained by setting the levels by setting one level for the background part and equally dividing the other part into 15 levels.

The Huffman coding means (309) performs Huffman coding, which makes compression by means of variable length code assignment of the gradation information output by the one gradation block approximation coding means (306), two gradation block approximation coding means (307) and linear quantization part (308).

Also, the MMR coding means (310) assigns 0 to all of the picture elements of the output from the label ① one gradation block approximation coding means (306), and assigns 1 to all of the picture elements of the output from the label ③ linear quantization part (308) based on the resolving power information obtained by the label ② two gradation block approximation coding means (307), and MMR-codes all these.

The code assembly part (311) generates code output based on the output of the Huffman coding means (309) and the MMR coding means (310). Figure 4 illustrates the code structure assembled to store one character region in the memory (25). It codes the following information.

Header information H: Image size in the character region and location of the region

Label ①: Gradation information G

Label ②: Gradation information G and resolving power information P

Label ③: Gradation information G

The character image coded in this manner is stored in the memory (25).

Next, the decoding of information on the character region that is coded and stored is performed by reproducing the concentration distribution of each block from the gradation information and resolving power information. Each of the blocks of label ① and label ② is reproduced in one gradation and two gradation (representative values). With regard to label ③, 16 representative values reproduce the concentration value of each picture element in the block.

Next, we will explain the processing of the halftone region.

Figure 5 illustrates an example of the constitution of the halftone region coding means (14) in Figure 1. It comprises a sampling part (51), predictive coding means (52), curved surface approximation means (53), approximation error computation part (54), nonlinear quantization part (55), Huffman coding means (56), mode determination means (57) and code assembly part (58).

According to the algorithm of the coding in this example, first, the sampling part (51) takes samples in a grid form to use the correlation between concentration values in vertical,

horizontal and diagonal directions. As this sampling position is fixed, no information that expresses the position of the sample point is necessary.

Next, as the variance of concentration values between adjacent sample points is small, the predictive coding means (52) makes predictive coding of the concentration value of the sample point to reduce the amount of codes.

Then, the approximated region surrounded by four adjacent sample points is approximated by the curved surface approximation means (53) that uses congruent tertiary batch [compression]. The mean square error between the original image in the region and the approximated image is then obtained for each region by the approximation error computing means (54).

The mode determination means (57), if the mean square error is larger than the threshold value, will compression code the approximation error (additional information) of each picture element in the region by means of the nonlinear quantization part (55) and the Huffman coding means (56).

Next, each process will be explained in more detail.

The sampling part (51) takes samples in a grid form as illustrated in Figure 6 to use the correlations between concentration values in the vertical, horizontal and diagonal directions. With regard to the space between sample points, to minimize blur in a part where the concentration change at the time of approximation is severe, it will be four picture elements. The concentration value of the sample point is assumed to be the mean concentration value of the four picture elements indicated with the shaded area in Figure 6. The sample point is not located at the picture element but located at the boundary between one picture element and another.

As illustrated in Figure 7 the predictive coding means (52) assumes the mean value of the four reference sample points adjacent to the vicinity of the noted sample point to be the predicted value. The prediction error is obtained by subtracting this predicted value from the concentration value of the noted sample point. Linear quantization is applied to this prediction error, and an approximation is made with 19 representative values. Next, a Huffman code is assigned to each representative value, which will be coded.

The curved surface approximation means (53) performs the process of approximating the region (the diamond-shaped region with the four vertices as the vertices, hereafter, called the approximated region) surrounded by four sample points applying congruent tertiary batch. The congruent tertiary batch interpolates the approximated region in reference to the 16 sample points illustrated in Figure 8.

The approximation error computing means (54) obtains for each region the mean square error (MSE) between the original image and the approximated image within the region as illustrated in Figure 9 to examine how much errors there are within the approximated region.

The mode determination means (57) determines that the mode is 1 (region that needs additional information) if the mean square error (MSE) is larger than the threshold value (TH), and that the mode is 2 (region that needs no additional information) if it is less than the threshold value (TH). Here, the mode information that prepares the mode information (MD) that indicates which mode the approximated region belongs in is binary image of 0 and 1, so MMR coding is adopted at the time of coding.

When the determination by the mode determination means (57) is mode 1, the approximation error is coded by means of the nonlinear quantization part (55) and the Huffman coding means (56). In this application example, the nonlinear quantization part (55) approximates with 11 representative values. Then, the Huffman coding means (56) codes by assigning a Huffman code to each representative value.

As illustrated in Figure 10 the code assembly part (58) assembles in code form the predicted error (S) of the sample point, mode information (M), and the approximation error, that is, additional information (F) of each picture element within the approximated region into coded output. The header information (H) is inserted at the beginning of the coded output. The header information is information such as the position and size of the region when the image in the halftone region is coded. The additional information is only needed in case of mode 1.

As explained above, the decoding of the image of the coded halftone region can be executed by a decoding means, which has, for instance, a constitution such as illustrated in Figure 11. From the code information of the code structure illustrated in Figure 10, the information separating means (111) separates the prediction error information (S), additional information (F), mode information (M), etc., and decodes these in respective decoding means (112), (114), and (117). Based on the concentration information of the sample point output from the decoding means (112) of the predictive coding code, the curved surface approximation means (113) interpolates by means of congruent tertiary batch, and obtains the concentration value of each picture element. When the mode information is mode 2, the selection means (116) will assume the output of the curved surface approximation means (113) with no modification to be the decoded output and make selection output. When the mode information is mode 1, the correction means (115) will correct the output of the curved surface approximation means (113) with the approximation error value, which has been decoded by the additional information decoding means (114), and the selection means (116) will select and output the output of the correction means (116) [sic; (115)] as the decoded output.

In the aforementioned application example (the first application example), we explained a case where the batch size of the halftone region is fixed, that is, when the sample points are fixed with a certain space. However, it is also possible to make the sample points variable and change the batch size.

Next, as the second application example, let us explain the system of variable batch size (and therefore sample points).

In this system, the image to be coded is divided into many diamonds (parallelogram-shaped), that is, batches. The points that are needed to constitute each batch are extracted as the sample points. The concentration information of these sample points is stored. In decoding, the concentration value within the batch is kept based on the concentration information of this sample point. By executing this process for all the batches, the image is decoded. Determining the batch size is based on the fact that a halftone image comprises both parts that have severe concentration changes such as the surface of an object and regions with monotone concentration distribution such as the background. Efficient coding can be realized if fine sampling is applied to the regions with severe concentration changes and rough sampling is applied to monotone regions.

The point where the second application example structurally (functionally) differs from the first application example illustrated in Figure 5 is mainly that selecting the sample points in the sampling means (51) is based on the batch size. The other parts of constitution are approximately the same.

The batch is a diamond-shaped region with four sample points as vertices such as illustrated in Figure 12. It is a square inclined by  $45^\circ$  to the arrangement of picture elements. The sample points are picked at the center of four picture elements in virtually the same manner as the first application example. The mean value of these four picture elements is assumed to be the concentration value of the sample point.

The minimum batch size is  $m = 2$ . Corresponding to the changes of concentration, as illustrated in Figure 12, in a state where the point that is located at the top of the batch is fixed, the batch is expanded in a shape of hanging bell, and  $n$  is determined within the range of 0 to 7.

The procedure for the other determination is as follows:

- I) As illustrated in Figure 13, the concentration slope information in the state of minimum batch size ( $n = 0$ ) is obtained as reference information.
- II) The batch with  $n = 1$  is divided into minimum batch size, and the concentration slope is obtained in the same manner as reference information for each of the minimum batches.
- III) If the absolute value of the difference between the absolute value of concentration slope and the absolute value of reference information is larger than the threshold value (TH1),  $n = 0$  will be assumed to be the batch size. If it is at or less than the threshold value, the same process as mentioned above will be executed with  $n = 2$ . However, if the minimum batch where the batch already hit is included, it is assumed that  $n = 0$  irrespective of the concentration slope.
- IV) Varying  $n$  from 0 to 7, the batch size is determined.

Once the batch size is determined, the sample point can be extracted. Predictive coding is done for each sample point. Predictive coding can be done in the same manner as explained for the first application example. However, unlike the first application example, sampling is done while changing the batch size. Thus, there are two issues. That is, ① four points to be used for prediction do not always exist, and ② there may be cases where no points exist.

Considering these issues, we will discuss below how to deal with each of the issues. With regard to the problem ①, the mean value is obtained by using only the existing point. With regard to the problem ②, the value of the vertex at the top of the batch is used as the predicted value. In case of the top level on the screen, no value exists even for the vertex located at the top. Thus, the median of usable concentration value width is assumed to be the predicted value.

In the coding structure, information on the coordinates of the top vertex of the batch and the batch size will further be added to the structure in Figure 10.

#### Effect of the invention

According to the present invention, the document image as the object is classified into two regions of a halftone region and a character region, and the coding system that is appropriate to the statistical characteristics of each region is used. Therefore, advanced compression while maintaining high picture quality can be realized.

Also, according to the present invention, as it is constituted that in the coding of the halftone region, the region surrounded by four sampled points is approximated by applying congruent tertiary batch, and when the approximation error is larger than a fixed threshold value, the error information of the approximated image is added, the improvement of picture quality can be realized.

Furthermore, in the mode of the present invention where, in the coding of the halftone region, fine sampling is done for the aforementioned sampling points in a region with severe concentration changes of the image and rough sampling is done in a monotone region, more efficient coding will be realized.

#### Brief description of the figures

Figure 1 is a block diagram that illustrates the constitution of the coding means of the image processing device in an application example of the present invention.

Figure 2 is a diagram that illustrates the schematic constitution of an image processing device wherein the present invention is applied to a digital copier.

Figure 3 is a diagram that illustrates an example of constitution of the character region coding means in Figure 1.

Figure 4 is a diagram that illustrates the code constitution of the codes that the character region coding means outputs.

Figure 5 is a diagram that illustrates an example of constitution of the halftone region coding means in Figure 1.

Figure 6 is a diagram that illustrates the sampling points of the fixed density sampling.

Figure 7 is a diagram that illustrates the reference sample points for predictive coding.

Figure 8 is a diagram that illustrates the coordinate system of the congruent tertiary batch.

Figure 9 is a diagram to explain the mean square errors between original image and approximated image within a region.

Figure 10 is a diagram that illustrates the code structure in the halftone region.

Figure 11 is a diagram that illustrates the constitution to decode the halftone region.

Figure 12 is a diagram to explain how to determine the batch size.

Figure 13 is a diagram to explain the concentration slope.

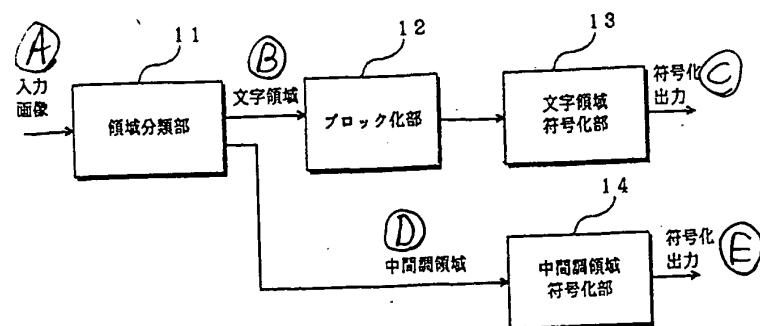


Figure 1

Key:	A	Input image
	B	Character region
	C	Coding output
	D	Halftone region
	E	Coding output
	11	Region classification means
	12	Block-making means
	13	Character region coding means
	14	Halftone region coding means

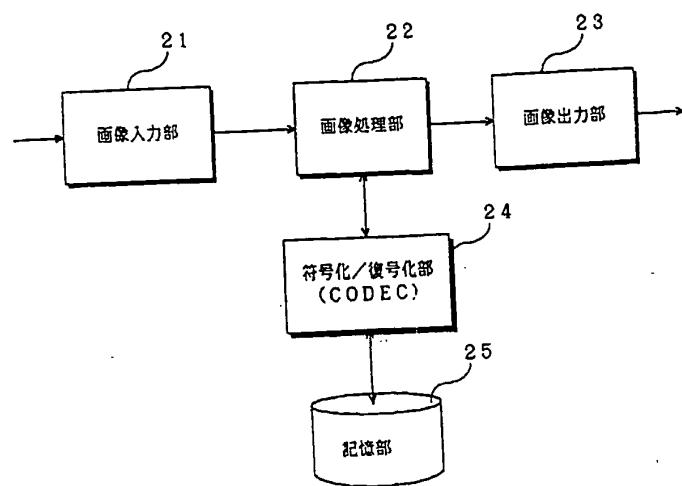


Figure 2

Key:

- 21 Image input part
- 22 Image processing part
- 23 Image output part
- 24 Coding/decoding (CODEC)
- 25 Memory

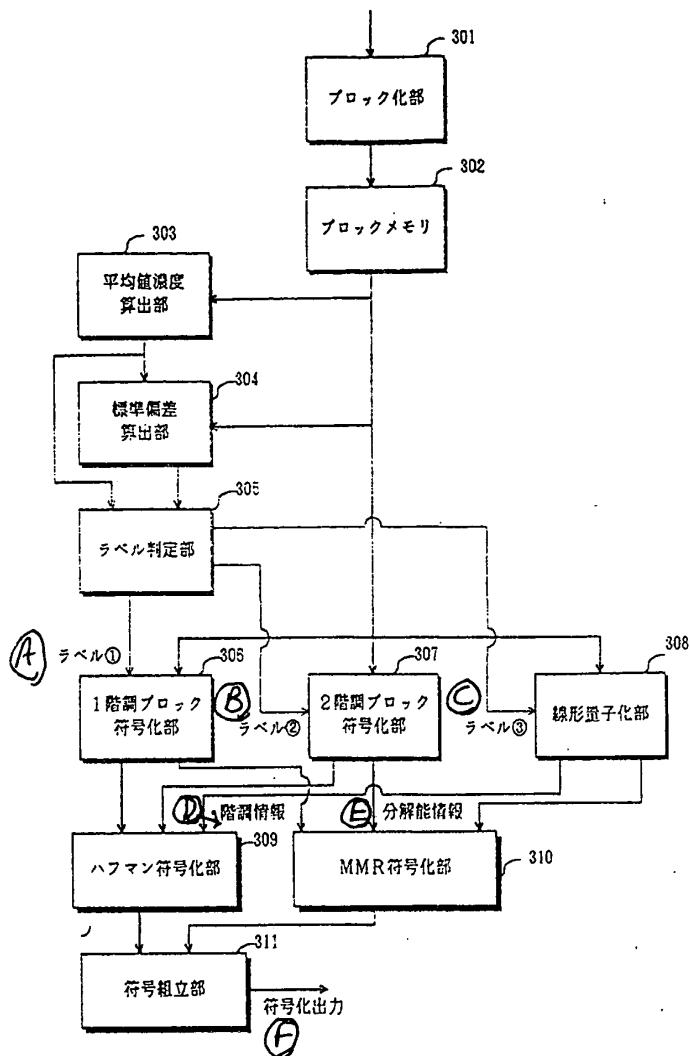


Figure 3

Key:	A	Label ①
	B	Label ②
	C	Label ③
	D	Gradation information
	E	Classification information
	F	Coding output
	301	Block-making means
	302	Block memory
	303	Mean concentration computing means
	304	Standard deviation computing means
	305	Label determination means

- 306 One gradation block coding means
- 307 Two gradation block coding means
- 308 Linear quantization means
- 309 Huffman coding means
- 310 MMR coding means
- 311 Code assembly means

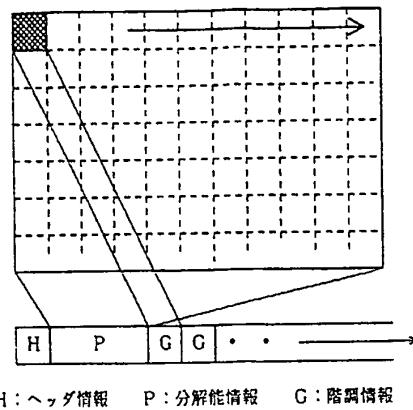


Figure 4

Key:

- G Gradation information
- H Header information
- P Resolving power information

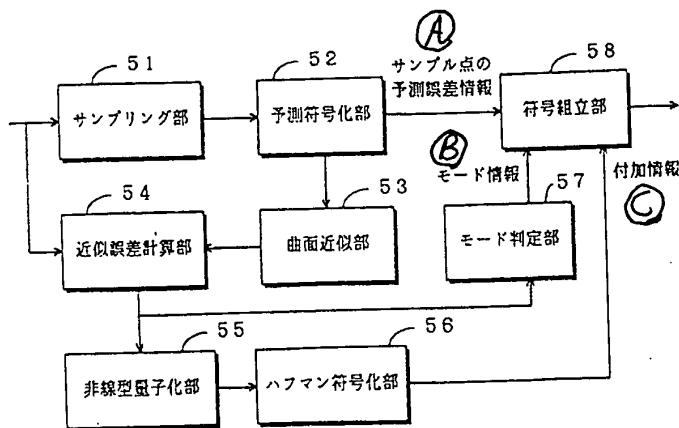


Figure 5

Key:

- A Prediction error information for sample point
- B Mode information
- C Additional information

- 51 Sampling means
- 52 Predictive coding means
- 58 Code assembly means
- 54 Approximation error computing means
- 53 Curved surface approximation means
- 57 Mode determination means
- 55 Nonlinear quantization means
- 56 Huffman coding means

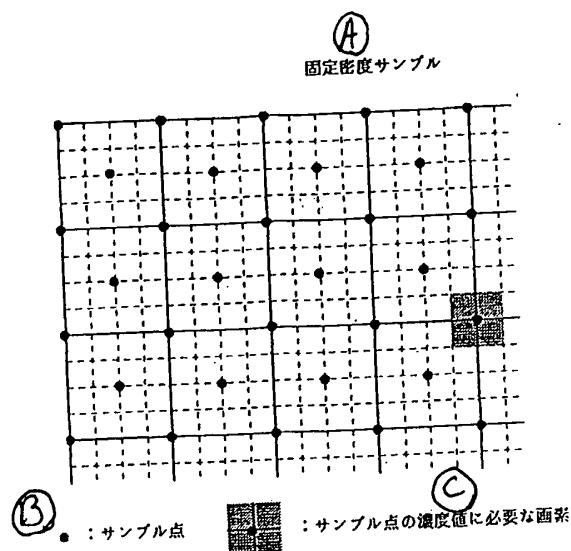


Figure 6

Key:

- A Fixed density sample
- B Sample point
- C Picture element needed for concentration value of sample point

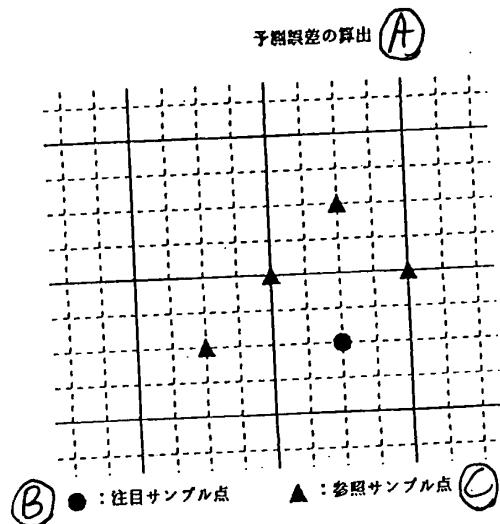


Figure 7

Key:

- A Computation of prediction error
- B Noted sample point
- C Reference sample point

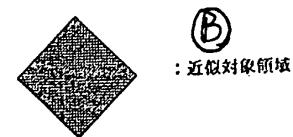
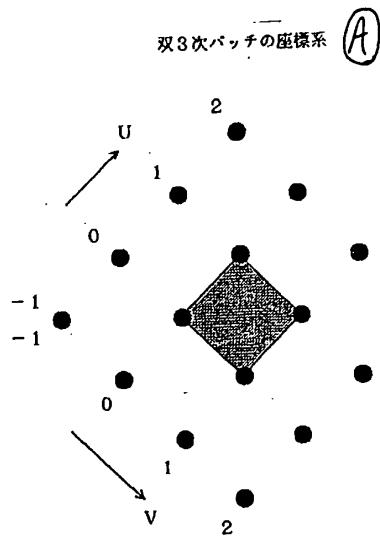
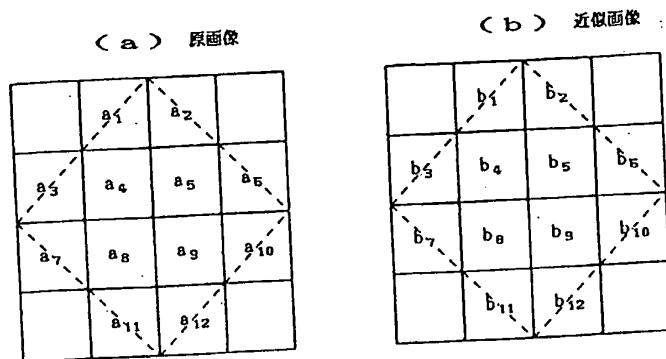


Figure 8

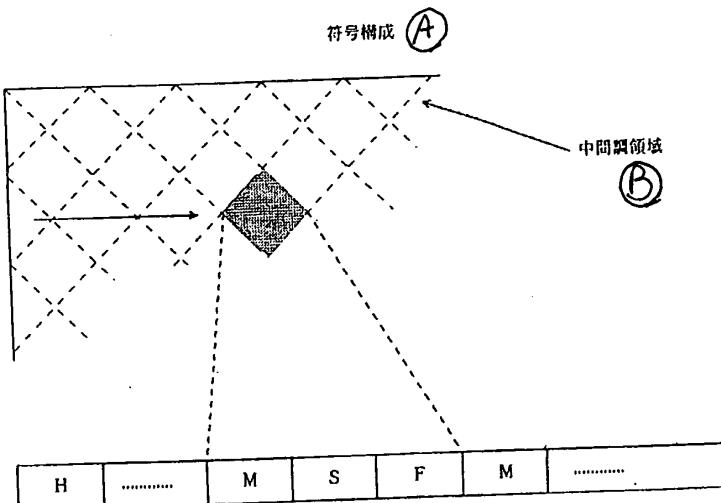
Key: A Coordinate system for congruent tertiary batch  
 B Approximated region



$$\textcircled{1} \quad \text{平均2乗誤差MSE} = \frac{1}{12} \sum_{i=1}^{12} (b_i - a_i)^2$$

Figure 9

Key: a Original image  
 b Approximated image  
 1 Mean square error MSE =



H : ヘッジ情報      S : サンプル点の予測誤差  
 M : モード情報      F : 付加情報 (モード1のみ)

Figure 10

Key: A Code constitution

B Halftone region  
 F Additional information (only mode 1)  
 H Header information  
 M Mode information  
 S Prediction error for sample point

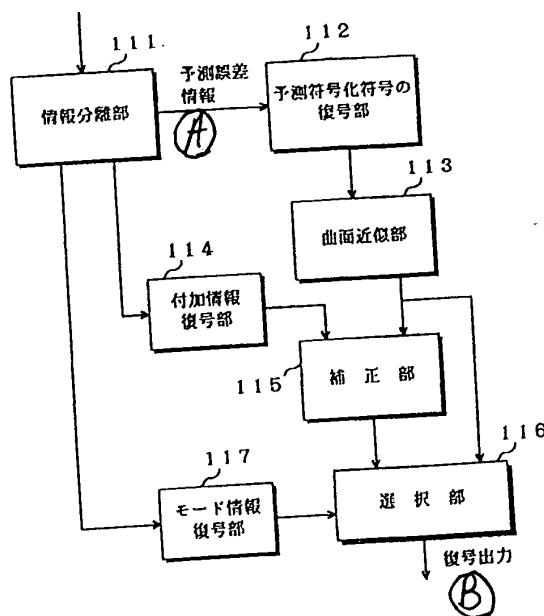


Figure 11

Key: A Prediction error information  
 B Decoded output  
 111 Information separating means  
 112 Decoding means for predictive coding code  
 113 Curved surface approximation means  
 114 Additional information decoding means  
 115 Correction means  
 116 Selection means  
 117 Mode information decoding means

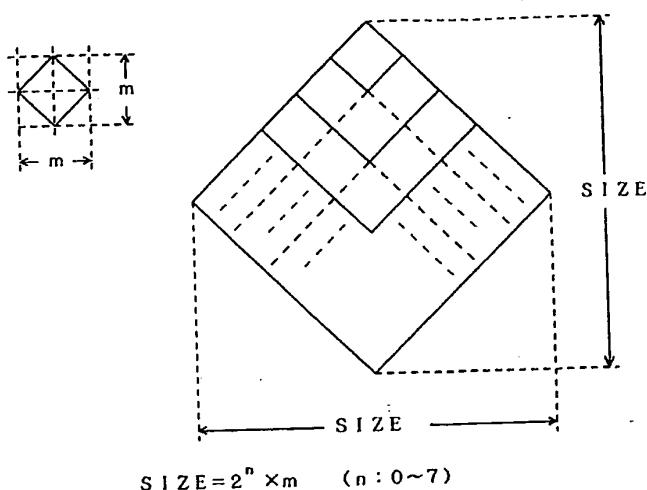


Figure 12. Size of the batch

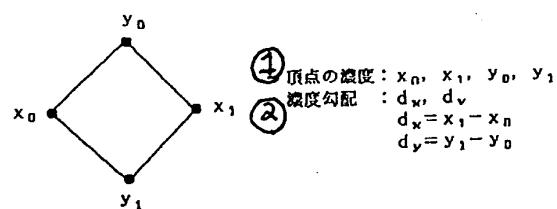


Figure 13

Key: 1 Concentration at the vertex  
 2 Concentration slope